ERIC Notebook

November 1999 Issue 7

UNC SCHOOL OF PUBLIC HEALTH
DEPARTMENT OF EPIDEMIOLOGY

Michel Ibrahim, MD, PhD
Director of Education Program

Lorraine Alexander, DrPH

Carl Shy, MD, DrPH

Sherry Farr, Graduate Research Assistant

ERIC
EPIDEMIOLOGIC RESEARCH & INFORMATION CENTER
Durham, North Carolina

Ron Horner, PhD
ERIC Director

The ERIC Notebook is funded by the Department of Veterans Affairs (DVA), Veterans Health Administration (VHA), Cooperative Studies Program (CSP), to promote the strategic growth of the epidemiologic capacity of the DVA.

Cross-sectional studies

Like cohort studies, cross-sectional studies conceptually begin with a population base. But unlike cohort studies, cross-sectional studies do not follow individuals over time, but only look at the prevalence of disease and/or exposure at one moment in time. These studies take a "snapshot" of the proportion of individuals in the population that are diseased and nondiseased at one point in time. Cross-sectional studies also differ from cohort studies in the populations that are studied. Cohort studies begin by selecting a population of persons who are at risk of disease, while cross-sectional studies begin by selecting a population group and then obtaining data to classify all individuals in the group as either diseased or nondiseased.

Study	Cohort	Cross-sectional
Study group	Population-at-risk	Entire population (or a sample)
Measures	Incidence and prevalence	Prevalence

Types of cross-sectional studies

There are two main types of cross-sectional studies. The first type, solely descriptive cross-sectional studies simply characterize the prevalence of disease in a specified population. Prevalence can be assessed at either a point in time (point prevalence) or over a defined period of time (period prevalence). Period prevalence is required when it takes time to accumulate sufficient information on disease in a population, e.g. what proportion of persons served by a public health clinic over a year have hypertension.

The second type may be called analytical cross-sectional studies in which data on the prevalence of exposure and disease are obtained for the purpose of comparing disease differences between exposed and nonexposed. Comparison of differences is the analytical component of these studies.

Analytical studies attempt to describe the prevalence of disease or non-disease by first beginning with a population base. These studies differ from solely descriptive cross-sectional studies in that they compare the proportion of exposed persons who are diseased (a / a+b) with the proportion of nonexposed persons who are diseased (c / c+d).

Calculating prevalence

The prevalence of a disease is simply the proportion of diseased individuals in a population.

- Prevalence = (cases) / (total population).

For the following example, two different measures of prevalence can be calculated: the prevalence of coronary heart disease (CHD) among the exposed (people who are not active) and the prevalence of CHD among the unexposed.

	Present CHD		Absent CHD		Total
Not active	50	a	b	200	250
Active	50	c	d	700	750
Total	100			900	1000

P1= a/a+b= 50/250 = 20% prevalence of CHD among people who are not active.
P0= c/c+d = 50/750 = 6.7% prevalence of CHD among people who are active.

The prevalence odds ratio

The prevalence odds ratio (POR) is calculated in the same manner as the odds ratio.

- POR = ad / bc

The POR estimates the incidence rate ratio (IRR) if the risk factor occurs over an extended period of time and if the duration of the outcome is not affected by exposure status. In the preceding example of activity level and CHD the POR equals (50 x 700) / (200 x 50) or 3.5. The POR in this study estimates the IRR and is interpreted to mean that

the estimated incidence rate among the exposed is 3.5 times greater than that among the unexposed.

The prevalence ratio

The prevalence ratio (PR) is analogous to the cumulative incidence ratio (CIR) of cohort studies. The denominators for both ratios are fixed populations -- fixed at the start of the study in the case of a cohort study, and fixed at the point or period of time for the case-control study. The prevalence ratio is calculated when the outcome occurs over a short period of time. For example, one would calculate a prevalence ratio for an acute outbreak of tuberculosis in a prison population. This is in contrast to calculating the overall prevalence of positive tuberculin skin tests among the prisoners.

The prevalence ratio can also be calculated from the information on CHD and physical activity. It is preferable to calculate the prevalence odds ratio when the period for being at risk of developing the outcome extends over a considerable time (months to years) as it does in this example.

- $PR = (a/N_1) / (c/N_0)$
- $PR= (50/250) / (50/750) = 3.0$

In this case, a prevalence ratio of 3.0 can be interpreted to mean that the proportion of people with CHD is 3-fold greater if a person is not physically active.

POR vs. PR

For chronic disease studies or studies of long-lasting risk factors, POR is the preferred measure of association in cross sectional studies. For acute disease studies, PR is the preferred measure of association. If the prevalence of disease is low, i.e. 10% or less in exposed and nonexposed populations, POR = PR. Since cross-sectional studies are particularly useful for investigating chronic diseases (e.g. prevalence of AIDS) where the onset of disease is difficult to determine, or for studying long lasting risk factors (such as smoking, hypertension, and high fat diets), the prevalence odds ratio will generally be the preferred measure of association.

POR and PR at a Glance	
POR	PR
Estimates the IRR	Estimates the CIR
Best for chronic diseases	Best for acute diseases

Limitations of Cross Sectional Studies to Evaluate Risk

Recall that, under steady conditions, the prevalence of disease is influenced both by incidence and duration of disease (or survival with disease).

- Prevalence = Incidence x Average Duration of Disease

Persons who survive longer with a disease will have a higher probability of being counted in the numerator of a prevalence proportion. Short-term survivors will be less likely to be counted as a case. Incidence is influenced only by exposure, whereas prevalence is influenced both by exposure and duration of disease.

If exposure influences survival time, then the POR or PR will not provide a valid estimate of the IRR or CIR. Thus, the interpretation of the POR or PR is subject to survival bias.

Even if incidence remains constant, either an improvement in disease treatment-that results in higher cure rates- or increased lethality resulting in a higher case fatality rate, will result in decreased prevalence. The disease itself or the threat of developing the disease may cause outmigration of cases from an environment perceived as causing disease, e.g. workers affected by toxic exposures in a plant may quit, while more resistant workers will stay. This selective migration can bias measures of prevalence.

Other Problems with Interpretation of Cross Sectional Studies

Cross-sectional studies as well as case-control studies are affected by the antecedent-consequent bias. This bias occurs when it cannot be determined that exposure preceded disease, since both are ascertained at the same time (unlike cohort studies or clinical trials). Antecedent-consequent bias does not affect cohort studies because subjects in cohort studies are selected for study because they are disease-free. Exposure is actually observed to precede disease only in a cohort design, including randomized trials.

Uses of Descriptive Studies

Descriptive studies are an important method to evaluate the proportion of a population with disease or with risk factors for disease, such as the prevalence of asthma in children or the prevalence of elevated blood lead in toddlers.

Descriptive cross-sectional studies are widely used to estimate the occurrence of risk factors in segments of the population characterized by age, sex, race or SES. National examples of cross-sectional studies of great importance are the decinnial census, the National Health and Nutrition Surveys (NHANES) and determining the prevalence of HIV positive antibodies in military recruits. Opinion polls and political polls are basically cross-

sectional studies. Surveillance of changes in smoking habits or of other behavioral risk factors are sequential cross-sectional studies. Similarly, surveillance of long lasting diseases such as AIDS are cross-sectional. Descriptive cross-sectional studies are useful for planning or administering preventive or health care services, surveillance programs, and surveys and polls.

Uses of Descriptive/Analytical Studies

Descriptive/analytical studies are useful for studying the association between exposure and disease onset for chronic diseases where researchers lack information on time of onset, such as diet and arthritis, smoking and chronic bronchitis, lead-induced hypertension, and asthma and exposure to allergens. Interpretation requires caution regarding potential association of duration of disease with exposure status (survival bias).

Survival bias may be minimized if information can be obtained on exposures that clearly preceded the first symptoms of a chronic disease such as arthritis, diabetes, or chronic bronchitis. The ability to accomplish this depends on access to medical records documenting the initial patient visits or examinations for the chronic disease, and possibly on historical records on the exposure of the individual prior to these first visits, e.g. where the person lived or where the person was employed.

Self-evaluation

Q 1: For which of the following disease situations could a cross-sectional study design be used to determine prevalence?
a. Disease w, a highly fatal disease in which the average length of survival after diagnosis is approximately one month.
b. Disease x, a disease with a long latency period where infection is identifiable by a serological test.
c. Disease y, a disease which is caused by a chromosomal malformation present at birth (or before).
d. Both b and c

Q 2: (Adapted from Norell S. Workbook of Epidemiology. New York: Oxford University Press, 1995.)

A hypothetical study of the effect of alcohol intake on the risk of gallstones in Mexican-Americans aged 20-74 was conducted. Ultrasonography of the gallbladder of 2200 study subjects identified 152 subjects with gallstones, 42 of whom listed their alcohol consumption level as "high". Of the men without gallstones, 220 listed their alcohol intake level as "high". Interviews were conducted at the same time as the examinations to determine the exposure level of study participants.

Which measure of association is most appropriate in this cross-sectional design?

a. Prevalence ratio
b. Prevalence odds ratio
c. Cumulative incidence
d. None of the above

Answers:
1. Correct answer: d. Both b and c are situations in which a cross-sectional study design could be used to determine a measure of prevalence. Prevalence can be determined if a disease has a long latency period that can be detected through a serological test. In choice c, the duration of disease is not affected by exposure. The condition is due to a chromosomal malformation, and will be present throughout life regardless of how long a person was exposed. Prevalence is often used to describe conditions that are permanent, and present at birth, such as congenital malformations.

Explanation of other choices:

a. Incorrect. In this situation, highly fatal cases would not be represented in a measure of prevalence since many people with the disease may have died before the questionnaire was administered. Prevalence is not an accurate measure when a disease is highly fatal.

2. Correct answer: b.
The prevalence odds ratio (POR) is the correct measure of association because POR is preferred in studies of chronic diseases or studies of long-lasting risk factors, such as alcohol intake.

Explanation of other choices:
a. Incorrect. The PR is calculated when the outcome occurs over a short time, such as an acute outbreak of an infectious disease. Gallstones develop more slowly, so the POR would be more appropriate.
c. Incorrect. The cumulative incidence (CI) is a measure of association that can only be calculated from a cohort study. It requires a count of total new cases of a disease divided by the population at risk. This is not possible from a cross-sectional study since subjects are not observed over time.
d. Incorrect.

Glossary

Antecedent-consequent bias: occurs in cross-sectional studies when it cannot be determined if exposure preceded disease.
Prevalence: the proportion of diseased individuals in a population.

ERIC NOTEBOOK IS PRODUCED BY THE EDUCATIONAL ARM (MICHEL A. IBRAHIM, MD, PHD, DIRECTOR) OF THE EPIDEMIOLOGIC RESEARCH AND INFORMATION CENTER AT DURHAM, NC (RON HORNER, PHD, DIRECTOR)

If you would like to receive ERIC Notebook please fill out the form below:

Name: _____

Degree(s):_____

Address:_____

City, State, Zip:_____

Telephone Number:_____

Fax Number:_____

E-mail Address:_____

Please fax to: 919-966-2089 – Attn: Sherry Farr

Mail to : UNC-CH , School of Public Health , Attn: Sherry Farr Department of Epidemiology , Campus Box 7400, Chapel Hill, NC 27599

Upcoming Topics

Please let Sherry Farr.know which topics are of special interest to you so that we can include them in a future issue.

UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL
SCHOOL OF PUBLIC HEALTH
DEPARTMENT OF EPIDEMIOLOGY
CAMPUS BOX 7400
CHAPEL HILL, NC 27599